VOLUME 51
No. 1

WHOLE NO. 227
1939

Psychological Monographs

EDITED BY
JOHN F. DASHIELL
UNIVERSITY OF NORTH CAROLINA

Differential Forecasts of Achievement and Their Use in Educational Counseling

By
RALPH ROBINSON WOLF, Jr.
DEPARTMENT OF PERSONNEL STUDY, YALE UNIVERSITY

A DISSERTATION PRESENTED FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY
IN YALE UNIVERSITY

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PUBLICATIONS OFFICE
THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO

Psychological Monographs

Differential [illegible] of Achieve-
ment and [illegible]
Educational Counseling

[illegible]

ACKNOWLEDGMENTS

In the preparation of the materials of this research valuable assistance was received from several members of the faculty and other individuals at Yale University. I am especially indebted to Dr. Albert B. Crawford, Director of the Department of Personnel Study, and Dr. Paul S. Burnham, Research Assistant in that Department for their invaluable advice and suggestions throughout the course of this investigation. I am also indebted to Professors Mark A. May and J. Warren Tilton of the Department of Education, who read the completed manuscript and made many helpful comments pertaining to the revision of the material.

Mention must also be made of the undergraduate assistants in the Department of Personnel Study who carefully carried out the greater part of the statistical calculations, and of my wife who performed the laborious task of typing and reading the manuscript and proof.

I wish also to express my very real gratitude to the Carnegie Foundation for the Advancement of Teaching and the Institute of Human Relations, whose joint contribution, through the Department of Psychology at Yale University, to the research budget of the Department of Personnel Study, has been of material aid in the furtherance of this investigation.

R. R. W.

NEW HAVEN, CONNECTICUT
April 15, 1937

TABLE OF CONTENTS

CHAPTER I

Statement of the Problem

The accurate prediction of individual trait differences in scholastic performance is one of the ultimate goals of a scientific educational psychology and personnel research program. Although this general problem is of major concern to educators at every school level, the present investigation is limited in scope to a study of trait differences among a sampling of the student body of Yale University.

The problem of predicting differences in academic performance may be approached from two major standpoints, characterized roughly as the admissions' point of view and that of guidance or counseling. While this study is concerned only with the latter aspect of the problem, it seems necessary at this point to elaborate upon the former in order to make explicit the distinction between the two. From the standpoint of a college admission board it is important to know the degree to which applicants to the college differ, relatively to each other, in general scholastic promise. In 1927 the newly founded Department of Personnel Study at Yale University investigated the possibility of combining prematriculation academic data (school record, entrance examination average, age, and scholastic aptitude test score) into a single index predictive of individual grade averages in the Freshman Year. General predictions of this type have been calculated for each entering group at Yale beginning with the Class of 1931. Subsequent comparison of these predicted grades with actual Freshman Year grade averages has consistently yielded correlation coefficients around +.70. These forecasts accurately predict Freshman grade averages within a range of plus or minus four points for 71% of the cases; 85% are accurately predicted within a range of plus or minus six points; and 94% of the cases are correctly forecast within a range

of plus or minus eight points. This type of prediction, general in nature, is designed to indicate which students possess superior capacity for college work as a whole; it offers a scientific basis for the administrative problem of selective admissions.

After a student has been admitted to Yale, he is faced with the necessity of selecting a course of study from among many curricular offerings, and, from the standpoint of counseling the individual with reference to his choice, a somewhat different problem from that of selective admissions is raised. In this instance it is of primary importance to measure the extent to which a given individual differs, within himself, with reference to his several abilities; it is only of secondary importance to know the extent to which his abilities are differentiated from those of other individuals.

According to the educational plan at Yale undergraduates are admitted to a common Freshman Year in which the course of study, though somewhat prescribed, is sufficiently flexible to permit a small range of selection within certain groups of subjects to meet the varying requirements of the three upper-class schools. In general, students desiring to study scientific subjects subsequently enter the Sheffield Scientific School; those planning to study engineering enter the School of Engineering; and those who aim to follow a so-called "academic" or "cultural" program enter Yale College. Although the major differences between the curricula of the upper-class schools are not clearly evident in the prerequisite work for these advanced divisions, important differentiation between individual programs of study actually begins in the Freshman Year and is none the less real.

At the present time the choice of an individual's course of study at Yale is determined by his expressed interests on the one hand and, on the other, by the subjective estimates of his capacity for various types of curricula made by faculty counselors and certain administrative officers of the University (in addition to the advice of his family, friends, secondary school teachers, and others). In an effort to improve upon these subjective judgments through the development of objectively determined indices of differential capacity, the Department of Personnel Study has

experimented for several years with the problem of calculating differential predictions from certain prematriculation data. Two such predictions have been made. One has been called a " verbal prediction " designed to forecast ability in highly verbal subjects, such as English, History, and the Foreign Languages; and the other has been termed a " quantitative prediction " designed to forecast ability in subjects involving quantities as content, such as Mathematics and the Sciences.

The frequent existence of real differences between individual capacity for verbal and quantitative types of curricula is one of the major assumptions of this study. Whereas its validity with reference to the data of this research will be examined in a subsequent chapter, it seems appropriate to mention at this point that the work of other investigators supports this contention and that the results of the preliminary experimental work of the Department of Personnel Study upon this problem tend, in a substantial number of cases, to indicate the existence of fundamental differences of the type just mentioned.

In addition to such experimental data there is, of course, a rational basis for postulating a difference between verbal and quantitative materials and for inferring differences in individual capacity for each type of work. These materials differ in their fundamental modes of expression; the methods and techniques for dealing with each type of material differs; and the results gained from their study appear to have basically different characteristics. In Mathematics and the Sciences (mainly Chemistry and Physics) full use is made of quantitative terminology; mathematical symbols, and other scientific and technological abbreviations common to formulae and equations, form both the language of and methods for dealing with these subjects. In contrast to this mode of expression, English, History, and the Foreign Languages utilize verbal symbols or words almost wholly, and in working with these materials the method of attack is descriptive, interpretive, and expository in nature. Furthermore, the results of the latter group of subjects may be said to be much more an expression of opinion or qualitative judgment, and are thus more debatable than the final outcome of Mathe-

matics and the Sciences, the results of which are expressed in relatively more objective, quantitative, and exacting terms.

Subsequent to a brief presentation in the next chapter of previous investigations at Yale and elsewhere related to the problem of predicting differential academic abilities of various types, it will be the writer's purpose to pursue the following problems:

(a) The differences between individual achievement in verbal and quantitative types of subject matter—the extent of their existence, and their prediction.

(b) The validity of these differential forecasts of achievement in verbal and quantitative subjects of study.

(c) The feasibility of utilizing these differential forecasts in educational counseling work at Yale—the limits of their application, and their interpretation.

CHAPTER II

Review of Previous Related Investigations

Research in the field of differential psychology falls naturally into two major categories: (a) the determination of *individual differences* or the extent to which a group of individuals differ, relatively to each other, in a single trait; and (b) the determination of *trait differences* or the extent to which a single individual differs, within himself, in two or more traits. In the first chapter the measurement of individual differences was referred to as the problem of selective admissions, and the measurement of trait differences was designated the problem of guidance or counseling. This study is concerned only with the latter, a category of research which has received relatively little attention from investigators as compared with the voluminous literature on differences between individuals. Furthermore, the following review deals only with *educational* or *scholastic traits* as distinguished from vocational or manual skills, and with the *prediction* of trait differences as distinguished from their present determination.

Kelley (20) was perhaps the first investigator to report on differential educational abilities. His data published in 1914, although by no means conclusive, indicate that differences in ability in English and Arithmetic, found at the elementary school level, tend to persist and appear four years later as differences of the same type in high school.

In 1920 Thorndike (27), without providing supporting evidence for his contention, outlined three types of intelligence which he thought were possessed by individuals: (a) abstract intelligence, or the ability to understand and manage ideas and symbols, such as words, numbers, chemical or physical formulae, legal decisions, scientific laws and principles, and the like; (b) social intelligence, or the ability to understand and manage people—to act wisely in human relations; and (c) mechanical intelligence, or the ability to understand and manage things and mechanisms. Thorndike expressed the belief that all three of

these types were possessed by most people in about the same amount, but that certain individuals were outstandingly superior or inferior in one or another of the three.

Kelley (21) in summarizing his later work in this field expressed the view that his findings, in the main, corroborated the conclusions published by Thorndike (28) in 1921. The following traits were found to be more or less independent of each other: (a) verbal intelligence, or the ability which in the main underlies facility in naming opposites, coördinates, subordinates, supra-ordinates, and predicates; (b) quantitative intelligence, or the ability in the main underlying facility in computation and other situations involving numbers as content; and (c) spatial intelligence, or the ability in the main underlying facility in handling form boards, geometrical forms, and other similar tests. This outline provides a somewhat different conception of the human intellect than that presented by Thorndike (27) in 1920.

In 1925 Hull and Limp (17 and 19) published the results of an investigation into the possibility of forecasting the differential aptitudes of 73 high school students in Shorthand, Typewriting, English, and Algebra. Predicted scores for each of these school subjects, secured by means of separate test batteries and regression equations, were found to correlate with their criteria (teachers' marks) to the extent of +.51, +.61, +.65, and +.74, respectively. In order to secure a measure of the differences between these various aptitudes, the predicted scores for successive pairs of subjects were subtracted, yielding a column of plus and minus differences. Next, the corresponding criterion scores were subtracted, yielding a second column of plus and minus differences. This technique, which has been called the " indirect method " for finding a difference, produced the following coefficients of correlation between predicted and actual differences for various paired combinations of subjects:

Typewriting–English	.50
Shorthand–English	.42
Shorthand–Algebra	.42
Typewriting–Algebra	.31
English–Algebra	.23
Shorthand–Typewriting	.17

From these data, Hull and Limp concluded that there probably are real differences in the aptitudes of individuals for such similar activities as learning the subjects taught in high school.

In connection with the preceding investigation, Limp (17 and 23) developed a more direct method for predicting a difference between the paired aptitudes of an individual. Utilizing only that portion of the test data pertaining to English and Typewriting, he was able to predict a difference in aptitude directly by means of a single test battery and special regression equation so constructed that if the forecast came out negative it would mean that the individual probably would excel in Typing, whereas if it came out positive it indicated that the person probably would excel in English. This direct method for finding a difference in aptitude eliminates the necessity for calculating separate predicted scores for each school subject since the difference between these scores is all that is desired.

Some interesting comments have been made by Hull (17 and 18) on the problem of variability in the amount of different traits possessed by the individual. From a study of 35 complete series of comparable test scores on 107 first-year high school students he concludes: (a) that the distribution of talent within an individual seems to approximate the shape of the normal curve in the same sense that distributions of individual differences do, and presumably with the same implications; (b) that the extent or range of trait differences possessed by the individual (differences within a single individual in two or more traits) is approximately 80 per cent as great as individual differences (differences between two or more individuals in a single trait); and (c) that the wide differences in magnitude of trait variability suggest that the possibilities for scientific guidance may be about twice as great for certain individuals as for others.

An investigation into the problem of predicting differences in success between pairs of college subject groups was reported by Segel (24) in 1932. In this study the existence of differentiable mental abilities as represented by ten college subject groups was first established. The multiple coefficients of correlation between

predicted differences and actual differences for each of these groups were as follows ($N=97$):

* Languages–Economics	.694
* Economics–History	.632
* Languages–Physical Science	.555
* History–Physical Science	.543
Biological Science–Physical Science	.480
Economics–Biological Science	.438
History–Biological Science	.432
Physical Science–English	.427
Biological Science–English	.400
Languages–Biological Science	.333

* For these four groups about 70 per cent of the cases, counting from the mean, were found to be correctly forecast, *i.e.*, for 70 per cent of the cases, the predicted differences evidenced an idiosyncrasy in the same direction as that indicated by the actual differences between criterion grades.

From the general regression equation Segel developed a formula to predict differences in success between college subject groups directly. Letting X_1, X_2, X_3, etc., denote individual scores on various predictive measures, and letting X_L, X_{Ec}, X_B, etc., denote individual marks in different school subjects, he sought to determine $X_L - X_{Ec}$, $X_L - X_B$, etc., or $\overline{X}_d$, the estimated true differences between success in these subjects. Segel points out that the derivation of the following formula for $\overline{X}_d$ is the critical point of his study:

$$\overline{X}_d = b_{d1 \cdot 23 \ldots n} X_1 + b_{d2 \cdot 13 \ldots n} X_2 \ldots + b_{dn \cdot 12 \ldots (n-1)} X_n + C.$$

Lee and Segel (22) subsequently reported a simpler method for the development of differential prediction equations. This method uses zero order correlation coefficients between the predictive factors and each of the measures between which a difference is to be predicted. It eliminates the necessity of finding differences for each item, and in a study using several variables the number of correlations that must be calculated is thereby reduced. The regression equation for the prediction of (a–b) from X is:

$$\overline{X}_d = r_{(a-b)x} \frac{\sigma_{(a-b)}}{\sigma_x} (X - M_x) + M_{(a-b)}$$

where $$r_{(a-b)x} = \frac{r_{ax}\sigma_a - r_{bx}\sigma_b}{\sqrt{\sigma_a^2 + \sigma_b^2 - 2r_{ab}\sigma_a\sigma_b}}$$

Using as predictive factors the scores made by 354 students on parts II, IV, and V of the widely used American Council Psychological Examination, Segel and Gerberich (26) found a multiple correlation coefficient of +.467 between predicted differences and actual differences in English and Mathematics, and a coefficient of +.582 between predicted differences and actual differences in Foreign Languages and Mathematics.

The correlation of differences for 100 junior college students, between predictive measures, secured by combining scores on various psychological, achievement, and interest tests, and differences between grades in various college subjects, was reported by Segel (25) as follows:

Criterion (Differences in Marks)	Predictive Factors	Multiple Correlation Coefficients
Languages–Mathematics and Science	Iowa English Literature Iowa Mathematics Iowa Science Strong Engineering Scale Strong Law Scale Strong Life Insurance Scale	.49
Languages–History	American Council Psych. Exam. Iowa English Literature Iowa History and Social Science Strong Engineering Scale Strong Personnel Management Scale	.58
English–Mathematics and Science	Iowa Mathematics Strong Engineering Scale Strong Law Scale Strong Purchasing Agent Scale	.40
English–History	Iowa Science Strong Engineering Scale Strong Purchasing Agent Scale	.48
Mathematics and Science–History	American Council Psych. Exam. Iowa Mathematics Iowa Science Strong Engineering Scale Strong Purchasing Agent Scale	.72

Multiple correlation coefficients showing the relationship of differences in achievement between parts of the Iowa High-School Content Examination and scores on the Strong Vocational Interest Blank were as follows:

Criterion (Differences Between Test Scores)	Predictive Factors	Multiple Correlation Coefficients
Iowa English Literature–Mathematics	a. Strong Engineering Scale b. Strong Medicine Scale c. Strong Life Insurance Scale d. Strong Purchasing Agent Scale e. Strong Personnel Mgt. Scale	.68
Iowa English Literature–Science	(As above)	.54
Iowa Mathematics–History and Social Science	(As above)	.61
Iowa Science–History and Social Science	(As above except "d")	.87

Since its inception, the Scholastic Aptitude Test of the College Entrance Examination Board has been subjected to thorough analysis by Brigham and Broyler (2, 3, 4, 5). Application of Spearman's tetrad difference technique to the examinations of 1926 and 1927 led to the discovery of two distinct group factors measured by the various sub-tests. At first no attempt was made to name or attach psychological significance to these factors, but later reports of the Commission on Scholastic Aptitude Tests refer to them as the "verbal" and "mathematical" factors. The following quotation from the 1929 Report (4) will help to clarify the distinction which was found: "The present Scholastic Aptitude Test seems quite definitely related to English, Latin, history, and the more general reading courses in college. . . . It is limited in scope in that it seems to deal only with verbal symbolism, a type of symbolic thinking which is of prime importance if the student is expected to derive benefit from lectures and textbooks, but which is not of paramount importance in mathematics and science. A supplementary test of thinking in mathematical symbolism might serve the double function of revealing students who might profitably continue in the scientific fields, and of exposing those students who proposed election of a scientific curriculum as the result of verbal disability rather than a high degree of mathematical and scientific aptitude." From 1930 to 1935 inclusive, the Scholastic Aptitude Test contained both a verbal and a mathematical section.

In 1932 Crawford and Burnham (14) reported the results of an investigation designed to determine the validity of various

College Entrance Board Examinations as forecasts of Freshman grades at Yale in corresponding subjects of study. They concluded that "College Board Examinations yield disappointingly low correlations with Freshman work, whether the latter is measured by the general average in all first-year subjects, or confined to the same fields as the examinations have been specifically intended to test." This same article reports the results of an attempt to develop two differential predictions: one (academic) designed to forecast grades in English and History, and the other (scientific) designed to forecast grades in Mathematics and Chemistry. [These results are likewise reported in two other sources (**10** and **12**).] School ranks, College Board marks in specific subjects of study, and scores on the verbal (S.A.T.) and mathematical (M.A.T.) sections of the College Board Scholastic Aptitude Test were used as predictive factors. The coefficients of correlation between various combinations of these measures and Freshman grades for 439 students in the Class of 1934 were as follows:

Average of English and History grades vs:	
(a) C.E.E.B. English and History averages	.29
(b) Academic predictions based only on secondary school rank and S.A.T. scores	.55
(c) Academic predictions based on (a) and (b) combined	.56
Average of Mathematics and Chemistry grades vs:	
(d) C.E.E.B. Mathematics and Chemistry averages	.29
(e) Scientific predictions based only on secondary school rank and M.A.T. scores	.66
(f) Scientific predictions based on (d) and (e) combined	.67

The following year, 1933, Crawford (**13**) reported that 27 per cent of the Freshman class had a difference of eight points or more (on a marking scale of one hundred) between their "academic" and "scientific" predictions. Mid-year grades in the contrasted fields substantiated these predictions for 82 per cent of the above cases. In other words, these differential forecasts proved correct four out of five times for that quarter of the class which evidenced differences of a significant order.

In 1934 Crawford (**9**) reported that Freshman Year grade averages in English and History for the Class of 1937 correlated

with academic predictions based on secondary school ranks and S.A.T. scores to the extent of +.65; Mathematics and Science averages correlated +.64 with scientific predictions based on school ranks and M.A.T. scores; and the average of grades in Freshman Science, Mechanical Drawing, and Mathematics correlated +.70 with engineering predictions based on a combination of McCauley's and Mann's Spatial Relations Tests, and Brigham's Mathematical Aptitude Test.

Continued work on the problem of developing differential predictions of achievement at Yale was reported by Crawford (**11**) in 1936. Correlation of predicted scores, for 495 members of the Class of 1938, with marks subsequently received in corresponding Freshman courses yielded coefficients of +.62 for the academic or verbal group and +.67 for the scientific or quantitative field. Differences in individual predictions correlated to the extent of +.53 with corresponding differences in criteria. Using as standards for evaluating the validity of these forecasts, either a relative difference between grades in the criterion groups of half the predicted difference, or an absolute difference of four points or more (equivalent to about half the standard deviation of the marks in question), Crawford found, for about one-third of the cases, that approximately 70 per cent validated their predictions, 5 per cent performed contrary to expectations, and 25 per cent had indeterminate records.

CHAPTER III

Description of the Data

The desirability of developing forecasts of differential achievement prior to the selection of subjects of study in the Freshman Year has already been discussed in the first chapter. Prematriculation data available for use in calculating these predictions are limited to scores on the verbal and mathematical sections of the Scholastic Aptitude Test, College Entrance Examination Board marks in specific subjects of study, and secondary school ranks in class. No other measures of individual academic competence are available for Yale applicants generally. The records of individuals in the Classes of 1934, 1935, and 1936 form the basis for this study.

A consideration of the reliability of the measures to be used in calculating differential forecasts of college achievement and the reliability of the criteria against which these forecasts will be validated subsequently is important for a clearer understanding of the work to follow. Unfortunately few data are available on the reliability of the College Board Examinations in specific subjects of study. These data pertain, for the most part, to the reliability of the reading and grading of the examinations in English, French, and Mathematics rather than to the reliability of student effort or achievement. Their inclusion here would be misleading since the reliabilities of these examinations vary from year to year and cannot be taken as measures of the reliability of other College Board examinations. This discussion deals, therefore, only with the verbal and mathematical sections of the Scholastic Aptitude Test, secondary school records, and college grades.

The reliability coefficients, as reported by Broyler (6), for the verbal sub-tests of the Scholastic Aptitude Test are given in

Table I. The figures in the column headed *r* are the correlations between the sum of the odd-numbered items and the sum of the even-numbered items answered correctly. The figures in the column headed $2r/(1+r)$ are the estimates of what the relia-

TABLE I

RELIABILITY ESTIMATES FOR VERBAL SUB-TESTS OF THE SCHOLASTIC APTITUDE TEST (Form H)

Sub-Tests	N	r	2r/(1+r)
Antonyms	386	.9286	.9630
Double Definitions	386	.8645	.9273
Paragraph Reading	386	.8729	.9321

bility would be for each entire sub-test. From these data and the intercorrelations of the three sub-tests, Broyler estimated the reliability of the total verbal score to be between +.97 and +.98. Because it is sometimes argued that the split-half technique gives results that are too high, Forms F and G of the verbal sub-tests were administered at one sitting to 92 graduate students. The reliability coefficient thus obtained was +.95.

Split-half coefficients of reliability for the mathematical section of the Scholastic Aptitude Test have never been published because it has been felt that, due to the highly speeded nature of this test, the results would be rather misleading. However, Findley (6, 7, 8) at Cooper Union Institute of Technology has reported the following coefficients of reliability between comparable forms of this test given a day apart:

TABLE II

RELIABILITY ESTIMATES FOR MATHEMATICAL APTITUDE TEST

	N	r
Form G vs. Form GA	757	.88
Form H vs. Form HA	826	.90
Form I vs. Form IA	615	.89

Before discussing the reliability of secondary school grades, it is necessary to explain the manner in which these data are treated at Yale (10). Instead of using the average of all of a student's grades, the significance of which varies greatly with the marking system, his rank in class is first ascertained. Then

the student is assigned an arbitrary mark corresponding, in a normal probability distribution, to the position he actually attained in his school. But school standards as well as their marking systems vary, and it is evident that any student's rank depends in part upon the competition, or average performance, of the whole group among whom the rank was attained. From test records and other data affording evidence as to the general level of ability characteristic of most of the schools sending students to Yale, it is possible to equate for variations in grading standards and thus obtain, for each applicant, an *adjusted school rank* which takes account both of his own rank in school and of the past relative rank of that school among others.

The possibility of finding the true coefficient of reliability for this measure of school achievement is remote due to the fact that the data, gathered from secondary schools all over the country, are not sufficiently complete. However, some estimate of the reliability of this measure can be derived from a knowledge of the validity coefficient between it and Freshman Year averages. This coefficient has been reported (**10**) as about +.60. Because the reliability of a measure cannot be lower than its validity (except by chance)*, it is reasonable to assume that the reliability coefficient of school grades is not likely to be lower than +.60. The upper limit of this coefficient is not known, but of course it cannot be greater than +1.00 and probably does not exceed +.85, the approximate reliability of college grades.

In the absence of more objective achievement testing at Yale, the only criteria available at the present time, against which to validate the differential predictions of college achievement to be developed subsequently, are grades in specific subjects of study. Although the unreliability of instructors' marks is generally recognized, the coefficients of correlation between first and second term grades for the Class of 1936, shown in Table III, are indeed reasonably high. In fact, they are higher than analogous meas-uses previously reported at Yale. Strictly speaking, these are not true reliability coefficients in the technical sense. Such

* Kelley, Truman L.: *Statistical Method,* Macmillan Company, New York City, 1924, p. 206.

discernible factors as changes in subject matter, changes of instructor, and variations in student attitude, interest, and motivation, in addition to ordinary chance factors, condition grades with varying effect and thereby tend to lower these coefficients. Even the most accurately determined grades for one term will not necessarily show close relationship with those

TABLE III

Coefficients of Reliability for Freshman Grades in Specific Subjects of Study

(Correlation between First and Second Term Grades for Class of 1936)

Course	N	r	
English	782	.841	
History	555	.833	
French	502	.813	
German	226	.840	
Latin	232	.872	
Average of verbal subjects			.840
Mathematics	507	.877	
Chemistry	451	.821	
Physics	60	.870	
Engineering Drawing	214	.891	
Average of quantitative subjects			.865
Average of total subjects			.851

assigned in the second term because of actual variations in student performance. Despite these variable factors, these correlations are the closest approximation to reliability coefficients obtainable.

The correlation between any two measures is limited by their respective reliabilities. In view of this consideration, a correlation of .840 for the verbal prediction and .865 for the quantitative prediction may be regarded as the maximum validity coefficients that could be secured consistently for any prematriculation indices of probable scholastic performance in Yale at the present time.

CHAPTER IV

Individual Idiosyncrasies in Educational Achievement and Their Prediction

The term " individual idiosyncrasy," as used in this study, will refer to differences between the verbal and quantitative abilities of a student as judged by comparing the average of his college grades in these two types of subject matter. For example, if a student is found to have an average of 85 in his verbal subjects and an average of 75 in his quantitative subjects of study, he is here considered to possess an educational idiosyncrasy. The terms " verbal " and " quantitative," whether used with reference to types of subject matter or to a student's abilities (either demonstrated or predicted) in these fields, are to be understood merely as descriptive categories. In the work to follow English, History, French, German, Latin, and Spanish will be designated the " verbal group " ; and Mathematics, Chemistry, Physics, and Engineering Drawing will comprise the " quantitative group." The average of students' grades in each of these groups will be known as the " verbal criterion " and the " quantitative criterion " respectively.

The study of individual idiosyncrasies among the student body at Yale is facilitated by a rather unique arrangement of the Freshman Year curriculum. Practically all Freshmen must elect at least one course, but not more than two courses, from each of the following categories:

I. English or History
II. Greek or Latin or a Modern Language or (in the case of prospective engineers) Engineering Drawing
III. Mathematics or a Natural Science

Because of these distribution requirements a substantial majority of student programs in the Freshman Year include course work in both the verbal and quantitative groups. This provides an opportunity to study the inter-relationships of achievement in

various types of subject matter, and to validate initially the forecasts of differential ability to be developed later.

In Table IV two sets of related data are presented for the Classes of 1934, 1935, and 1936 combined. The left-hand portion of the table shows the intercorrelation of grades for various Freshman subjects of study. These intercorrelations were first calculated for the three classes separately and then combined, in order to determine whether serious error would be involved in utilizing the coefficients for the combined classes as approximately representative of the coefficients obtained for each class taken separately. The variation between the coefficients of the three separate classes proved to be small in most instances, and it was found that the coefficients for the combined classes could serve as representative measures in each case without introducing a substantial error. The last two columns to the right of the table show the correlations of grades in each subject of study with the average of those grades, within the verbal and quantitative groups, which comprise the two criteria. In the work involving these latter correlations, self-correlation effect was eliminated in each instance by exclusion of those subject grades from the grade averages composing the criterion group against which the subject was being correlated.

The grade intercorrelations presented in the left-hand portion of the table were calculated for the purpose of discovering, if possible, an empirical basis for the postulate that within the verbal group and within the quantitative group respectively the subjects are more closely related to one another than they are to those subjects in the differentiated group. These data validate this postulate, but the degree of "specialized" common function within each group is not marked, and the differential relationship between the verbal and quantitative groups is not clearly defined. This being the case, one would not expect to find, for any given subject, substantial differences between the paired coefficients shown in the two columns at the extreme right of Table IV. These differences range from .08 points for Chemistry to .28 points for Engineering Drawing.

Although these criterion differences are not as large as one

TABLE IV

INTERCORRELATION OF FRESHMAN GRADES FOR VERBAL AND QUANTITATIVE GROUPS

(Classes of 1934, 1935, and 1936 combined)

		History		French		German		Latin		Spanish		Mathe-matics		Chemistry		Physics		Drawing		Verbal Criterion		Quantitative Criterion	
		N	r	N	r	N	r	N	r	N	r	N	r	N	r	N	r	N	r	N	r	N	r
Verbal Group	English	1634	.61	1628	.50	571	.44	913	.51	331	.35	1463	.39	1289	.49	211	.47	602	.24	2440	.58	1862	.43
	History			1178	.42	335	.39	672	.47	260	.41	844	.36	735	.49	165	.55	222	.12	1748	.57	1186	.44
	French					128	.53	703	.54	120	.49	917	.43	762	.46	163	.43	293	.16	1686	.55	1213	.44
	German							140	.49	13	*	396	.43	404	.47	43	.11	174	.25	609	.51	482	.42
	Latin									70	.54	321	.50	287	.55	106	.54	10	*	939	.61	540	.51
	Spanish											182	.32	152	.39	34	.37	81	.25	359	.47	244	.35
Quant. Group	Mathematics													1111	.68	111	.67	631	.52	1572	.48	1249	.70
	Chemistry															7	*	595	.42	1379	.58	1121	.66
	Physics																	17	*	225	.54	119	.66
	Drawing																			669	.26	643	.54
Verbal Criterion																						1978	.53

* Number of cases too small to yield significant coefficients.
All coefficients in this table are positive and were calculated by the product-moment method.

would wish in a study of individual idiosyncrasies, it is interesting to note that, however small, they are in the direction desired in every case. That is to say, the subjects in the verbal group correlate consistently higher with the verbal criterion than with the quantitative criterion, and those subjects in the quantitative group correlate consistently higher with the quantitative criterion than with the verbal criterion. Furthermore, the average of coefficients for the verbal and quantitative groups of subjects with their respective criteria (.59) is higher than the average for these subject groups with their opposing criteria (.45). Without attempting a careful factor analysis of these data, the results of this crude approach would seem to indicate the possible existence of unique factors within each group, but the comparatively high correlation of .53 between the verbal and quantitative criteria would also indicate the presence of a common or general factor to a fair degree.

Considering the relatively small degree of differentiation between the verbal and quantitative criterion groups, it may be pertinent to ask at this point whether it is profitable to proceed further with this investigation. It is quite possible that the subjects in the Freshman curriculum at Yale are hybrids calling for a mixture of abilities, or there may be no clearly defined idiosyncrasies of the type dealt with in this study. The question remains, however, whether or not the existence of differential abilities is obscured in the Freshman Year by virtue of the fact that students must conform to the requirements governing distribution of course work, and must satisfactorily complete *all* this work before progressing to their major field of concentration in the upper-class years.

Let us take as an example a student who appears to possess a high degree of mathematical and scientific intelligence, but who is notably weak in his ability to master verbal subject matter. Let us further suppose that this student fulfills the distribution requirements in the Freshman Year by electing a typical Freshman course composed of English and/or History, a Modern Language, Mathematics, and one or two Sciences. In this situation it is not difficult to imagine that the student might spend

considerably more time on those subjects that are difficult for him (English, History, and Modern Language) than on those subjects that come easy to him (Mathematics and Science) in order to maintain a satisfactory average in all of his studies. As a result the student's record at the end of the year may show an average of 78 in his verbal subjects, an average of 80 in his quantitative subjects, and a general average of 79. If these numerical averages were considered alone, one might be led to the conclusion that the student in question possessed no marked differential ability. But these averages are not measures of ability alone; they are measures of expended effort as well. Hence the total result represents an average of inequalities wherein individual idiosyncrasies in ability may be obscured by formal faculty requirements, and the interplay of motivational factors (e.g., standards of eligibility for scholarship aid).

As explained in the previous chapter, prematriculation data available for use in calculating predictions of differential achievement are limited at Yale to scores on the verbal and mathematical sections of the Scholastic Aptitude Test, College Entrance Examination Board marks in specific subjects of study and secondary school ranks in class. From Table V it will be seen that of the 2582 students in the Classes of 1934, 1935, and 1936, 316 had incomplete prematriculation records, 40 were repeating their first year work, and 2226 had both complete prematriculation and Freshman Year records. But of the latter group 506 lacked a suitable quantitative criterion; all of their work was taken in verbal subjects of study. Consequently, only those 1720

TABLE V

ANALYSIS OF STUDENT RECORDS IN FRESHMAN YEAR

	1934	1935	1936	Total
Students having complete records	509	582	629	1720
Students having complete records but lacking quantitative criteria	159	174	173	506
Sub-total	668	756	802	2226
Repeaters in Freshman Year	9	9	22	40
Students having incomplete prematriculation records	173	85	58	316
Total entering Class	850	850	882	2582

students having both verbal and quantitative criterion grades in the Freshman Year were used as a basis for determining the two prediction equations set forth in Table VI. The 40 repeaters in the Freshman Year represent duplicate cases. These were included as members of their original classes.

The basic data used in the calculation of the verbal and quantitative predictions and the resulting regression equations are shown in Table VI. It will be noted that "Adjusted School Rank" is the only variable common to both predictions, and the one receiving the greatest weight in both regression equations. Because of the obvious importance of school rank in predictions

TABLE VI

FOUR-VARIABLE MULTIPLE CORRELATION DATA

(Classes of 1934, 1935, and 1936 combined)

Verbal Prediction

Variables	Means	Standard Deviations	Number of Cases
1. Freshman Verbal Average	74.0547	7.3589	1720
2. Adjusted School Rank	75.1709	6.5057	1720
3. Scholastic Aptitude Test Score	521.2674	89.4175	1720
4. College Board Verbal Average	70.4965	6.6481	1720

Zero Order Coefficients

$r_{12}=.5488$ $r_{13}=.4058$ $r_{14}=.4202$

$r_{23}=.3356$ $r_{24}=.2396$ $r_{34}=.3614$

$R_{1(234)}=.6423$ $\sigma_{1.234}=5.6406$

Regression Equation (score form)

$$X_1=b_{12.34}X_2+b_{13.24}X_3+b_{14.23}X_4+K$$
$$=.4874\ X_2+.0139\ X_3+.2833\ X_4+10.1991$$

Quantitative Prediction

Variables	Means	Standard Deviations	Number of Cases
1. Freshman Quantitative Average	71.5395	10.8984	1720
2. Adjusted School Rank	75.1709	6.5057	1720
3. Mathematical Aptitude Test Score	562.6047	83.5472	1720
4. College Board Quantitative Average	76.7465	9.7387	1720

Zero Order Coefficients

$r_{12}=.5757$ $r_{13}=.4193$ $r_{14}=.3405$

$r_{23}=.2693$ $r_{24}=.1999$ $r_{34}=.4225$

$R_{1(234)}=.6515$ $\sigma_{1.234}=8.2675$

Regression Equation (score form)

$$X_1=b_{12.34}X_2+b_{13.24}X_3+b_{14.23}X_4+K$$
$$=.8128\ X_2+.0295\ X_3+.1655\ X_4-18.8577$$

of this kind, it would seem advisable to differentiate if possible between secondary school averages in verbal and quantitative subjects of study. This could be done quite readily if individual marks were comparable from subject to subject and school to school. But the wide variations in secondary school grading standards necessitate the application of a rather lengthy and elaborate statistical procedure to these data before individual marks can be equated. Time has not been available for this further experimentation.

CHAPTER V

The Validity of Differential Forecasts in Freshman Year

In the work to follow questions of methodology will receive some special attention because the statistical procedure followed, although not new, is not well standardized. From the two regression equations given in Table VI, verbal and quantitative predicted scores were calculated for those 1720 students having complete entrance data and both verbal and quantitative criterion grades in Freshman Year. These scores were then transmuted to the mean and standard deviation of their respective criteria. The differential predictions were found by subtracting the quantitative from the verbal scores; and the differential criteria against which they were validated were found, in like manner, by subtracting the corresponding Freshman grade averages in these two fields. The writer is aware that these subtractions were not made in accordance with theoretically ideal statistical procedure, since the means and standard deviations of the verbal and quantitative distributions are not quite the same. Consequently a verbal criterion grade or predicted score of 75, for example, is not exactly comparable in this sense with a quantitative grade or score of 75 even though they are regarded as equivalent in the marking system.

It would seem, therefore, that a statistical error may have been introduced by subtracting incomparable scores. In order to determine whether this was actually true, both criteria and both predictions for the 582 usable cases in the Class of 1935 were transmuted to standard scores, all having the same mean and standard deviation. The resulting plots of plus and minus differences between the differential forecasts and their criteria were almost identical to the results obtained by using untransmuted scores.

As evidenced by the size of the coefficients obtained by other investigators in this field (cf. Chapter II), a high degree of correlation between differences is seldom secured. The correlations between predicted and criterion differences for the Classes of 1934, 1935, and 1936, separately and combined, given in Table VII are no exception to this general finding. However, they

TABLE VII

DATA PERTAINING TO THE CORRELATION OF ACTUAL DIFFERENCES BETWEEN VERBAL AND QUANTITATIVE GRADE AVERAGES IN FRESHMAN YEAR (X VARIABLE) AND CORRESPONDING PREDICTED DIFFERENCES (Y VARIABLE)

	Mx	σx	My	σy	N	r
Class of 1934	+2.85	8.88	+2.75	6.91	509	+.51
Class of 1935	+2.37	10.13	+2.39	6.26	582	+.47
Class of 1936	+2.48	9.63	+2.56	6.84	629	+.49
Classes combined	+2.56	9.59	+2.56	6.66	1720	+.49

do compare favorably with the correlation of +.53 obtained by Crawford (11) in a similar study of differential abilities for the Class of 1938.

Statisticians have generally maintained that predicted scores, having no higher index of relationship with their criteria than the correlations cited in the above table, are of negligible value due to the large error of estimate involved. It is currently fashionable to utilize the coefficient of alienation to point out the inadequateness of many predictive factors, such as those employed in this investigation, to forecast achievement with much better than chance results. For example, if one applies the formula for the coefficient of alienation to the data on the Class of 1934, the resulting coefficient is .86. It is customary for investigators to say no more about such a result than that the forecasting efficiency of the predictions is only 14 per cent better than chance. A conclusion of this sort, while statistically correct, is often misleading to the reader, who is inclined to accept it as a final judgment on the relative weakness of the predictions. In any thoroughgoing analysis it is necessary to proceed further and to discuss the validity of the predictions both with reference to the original purpose for calculating them and with reference to their eventual use.

The differential predictions in this research were designed for use in educational counseling work with students at Yale. In such work it is important to realize that the differences between the verbal and quantitative abilities of various individuals, whether demonstrated or predicted, are not of equal significance. Only a relatively small proportion of students have large differences of sufficient reliability to warrant their use in educational counseling. In order to demonstrate the significance of this statement, data from the scatter diagrams of predicted versus criterion differences for the three classes combined were reorganized and colligated in the form presented in Table VIII.

Because of the rather complicated organization of this table, detailed explanation seems necessary. It will be noted first that the number and percentage of cases are classified vertically (first column to left of table) according to the degree of predicted difference. Thus one finds, in the upper left hand corner of the table, that 132 students (8% of the total usable cases) had a predicted difference of 12 points or more between their verbal and quantitative forecasts. In the next sub-group one finds that 258 students (15%) had a predicted difference of 10 points or more; in the next sub-group 473 students had a predicted difference of 8 points or more, and so on down to the last row where it is noted that 1415 students had a predicted difference of 2 or more points.

Secondly, the number and percentage of cases within each of the aforementioned sub-groups are tabulated horizontally (main body of the table to the right) according to the magnitude of actual differences between criterion grades in Freshman Year. And thirdly, the differential performance characterizing the sub-groups at each level is denoted by the terms "positive," "indeterminate," and "negative" (second column of table). A "positive" differential performance refers to an actual difference between criterion grades in the same (positive) direction as forecast, and by the amount indicated at the head of each column to the right of the table; a "negative" differential performance refers to an actual criterion difference contrary (negative) to the direction forecast, and by the amount indi-

TABLE VIII

Number and Percentage of Cases in the Classes of 1934, 1935, and 1936 Combined, Classified Vertically as to Degree of Predicted Differences and Tabulated Horizontally as to Magnitude of Actual Differences between Criterion Grades in Freshman Year

(Differences counted from the approximate Mean)

Degree of Predicted Differences for Each Sub-Group Indicated Below	Differential Performance* Characterizing Each Sub-Group	Magnitude of Actual Differences in Criterion Grades																	
		1 or more		2 or more		3 or more		4 or more		5 or more		6 or more		7 or more		8 or more		9 or more	
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
12 points or more	Positive	113	*86*	107	*81*	103	*78*	98	*74*	96	*73*	91	*69*	86	*65*	79	*60*	68	*52*
132 cases	Indeterminate	3	*2*	11	*8*	21	*16*	29	*22*	31	*23*	38	*29*	43	*33*	50	*38*	61	*46*
(8%)	Negative	16	*12*	14	*11*	8	*6*	5	*4*	5	*4*	3	*2*	3	*2*	3	*2*	3	*2*
10 points or more	Positive	218	*85*	211	*82*	202	*78*	190	*74*	181	*70*	170	*66*	159	*62*	146	*57*	124	*48*
258 cases	Indeterminate	6	*2*	18	*7*	35	*14*	52	*20*	62	*24*	76	*29*	89	*34*	103	*40*	125	*49*
(15%)	Negative	34	*13*	29	*11*	21	*8*	16	*6*	15	*6*	12	*5*	10	*4*	9	*3*	9	*3*
8 points or more	Positive	382	*81*	372	*79*	355	*75*	326	*69*	311	*66*	290	*61*	268	*57*	244	*52*	216	*46*
473 cases	Indeterminate	12	*2*	38	*8*	69	*15*	106	*22*	125	*26*	154	*33*	179	*38*	206	*43*	239	*50*
(27%)	Negative	79	*17*	63	*13*	49	*10*	41	*9*	37	*8*	29	*6*	26	*5*	23	*5*	18	*4*
6 points or more	Positive	552	*78*	537	*76*	509	*72*	465	*66*	441	*62*	403	*57*	371	*53*	336	*48*	293	*41*
706 cases	Indeterminate	22	*3*	62	*9*	110	*16*	165	*23*	202	*29*	253	*36*	291	*41*	333	*47*	386	*55*
(41%)	Negative	132	*19*	107	*15*	87	*12*	76	*11*	63	*9*	50	*7*	44	*6*	37	*5*	27	*4*
4 points or more	Positive	722	*72*	703	*70*	667	*67*	604	*61*	572	*57*	516	*52*	469	*47*	424	*42*	371	*37*
998 cases	Indeterminate	49	*5*	105	*11*	179	*18*	263	*26*	316	*32*	396	*40*	451	*45*	514	*52*	580	*58*
(58%)	Negative	227	*23*	190	*19*	152	*15*	131	*13*	110	*11*	86	*8*	78	*8*	60	*6*	47	*5*
2 points or more	Positive	974	*69*	940	*66*	883	*62*	795	*56*	745	*53*	666	*47*	602	*43*	535	*38*	461	*33*
1415 cases	Indeterminate	64	*4*	157	*11*	276	*20*	402	*29*	481	*34*	596	*42*	680	*48*	769	*54*	861	*61*
(82%)	Negative	377	*27*	318	*23*	256	*18*	218	*15*	189	*13*	153	*11*	133	*9*	111	*8*	93	*6*

* Positive=in the direction forecast, by the amounts indicated.
Indeterminate=less than the amounts indicated in either a positive or negative direction.
Negative=contrary to direction forecast, by the amounts indicated.

cated; and an "indeterminate" performance refers to an actual difference between criterion grades of less than the amounts indicated in either a positive or negative direction.

In order to make this clear, let us consider that sub-group of 706 students having a predicted difference of six or more points. In the first column under "magnitude of actual differences in criterion grades" one finds that 552 students, or 78% of 706, had an actual difference of one or more points in the direction forecast (positive); 132 (19%) had an actual difference of one or more points contrary to the direction forecast (negative); and 22 students (3%) had a criterion difference of less than plus or minus one (i.e. zero) and were therefore designated indeterminate.

Before conclusions can be drawn from an analysis of these data, it is necessary to establish some degree of differential performance as a satisfactory criterion of validity. In view of the tendency for curriculum requirements to conceal, rather than to reveal, educational aptitudes in the Freshman Year, Crawford (11) adjudged either an *absolute* criterion difference of four points (representing approximately one-half the standard deviation of the criterion differences), or a *relative* standard, proportionate to one-half the difference forecast, as sufficiently rigid criteria for the purposes of his study. Needless to say, these criteria are more or less arbitrary determinations which have no necessary relationship to the accepted standards of statistical reliability. (See footnote on next page.) The reader who regards them as insufficiently rigorous is at liberty to evaluate the foregoing table in any way he sees fit. However, in the light of previous counseling experience, the writer is inclined to accept Crawford's criteria as adequate for practical purposes at this point in the present study.

Using a criterion difference of one-half the difference forecast as a standard in evaluating Table VIII, one notes a striking similarity between the proportion of cases correctly forecast (positive) at each level. Along this diagonal running through a series of coordinates from "a predicted difference of two or more with a criterion difference of one or more" to "a predicted

difference of twelve or more with a criterion difference of six or more," the proportion of cases correctly forecast does not vary by more than three per cent. Turning now to the other standard of validity—an absolute criterion difference of four points or more—one will note in this vertical column, as in all others, that the percentage of cases performing positively (in the direction forecast) increases gradually with the size of the predicted difference.

In summarizing these data it is simplest perhaps to apply both standards of validity suggested above. A predicted difference of eight points and a criterion difference of four points satisfy both these tests.* Consequently it may be stated that significant differences in performance have been correctly forecast in approximately 70% of those cases for whom the predicted difference exceeded seven points. Approximately 10% clearly contradict their forecasts, and the remainder (20%) fall into an indeterminate category. In other words, about one-quarter of the cases had eight or more points variation between their verbal and quantitative predictions. For this minority group, it is clearly evident (if the reader will grant the efficacy of the criteria of validity herein utilized) that the differential predictions developed in this study *correctly forecast differences in ability approximately two out of every three times in the Freshman Year.* This situation holds true as well for each of the three classes taken separately.

By means of Table VIII an attempt has been made, not only to make explicit the relative validity of the differential forecasts at various levels (degrees) of predicted difference, but also to make clear the fallacy of judging the predictive efficiency of one's data on the basis of total class results. The coefficients

* It was not possible to calculate accurately the reliability coefficients for predicted or criterion differences because of the incompleteness of the data. The following coefficients, therefore, represent approximations to the true results:

Estimated reliability of predicted differences=.25. $\sigma_{(\text{Meas.})}=5.78$
Estimated reliability of criterion differences=.68. $\sigma_{(\text{Meas}).}=5.40$

Hence, there are approximately 9 chances out of 10 that a predicted difference of 8 is a true difference; and there are approximately 3 chances out of 4 that a criterion difference of 4 is reliable.

set forth in Table VII are based on mass correlations. Analysis of them alone tends to obscure the significance of differences for *that critical minority group upon which counseling would ordinarily be based.* Theoretically, one would not expect more than a relatively small percentage of any class to evince marked differences between their aptitudes or achievements in disparate fields of study. Actually, most individuals are found to have no marked educational idiosyncrasies, and for them differential educational counseling is relatively unimportant. But for the critical minority of cases with large differences between their verbal and quantitative abilities, differential counseling may be justified.

Counselors are frequently asked to interpret a student's record and to suggest the relative chances for success in various fields of major study. In a practical guidance situation of this sort, the wise counselor will utilize such data as have been presented in the foregoing pages, or any other manifestations of individual idiosyncrasies, in proportion to the extent of variation which they indicate. It is evident that these data provide only imperfect tools for guidance work, and from this standpoint are not wholly satisfactory. But they are, nevertheless, an improvement over mere subjective estimates of ability, and, when placed in the hands of intelligent counselors, these differential forecasts may be used to advantage. However, it must be clearly borne in mind that the predictions herein developed have been validated only against Freshman Year records. The ultimate amount of dependence which may be placed upon them as practical guidance instruments cannot be vouchsafed until their validity in the upper-class years has been demonstrated. If, in this further work, a low degree of relationship is found between achievement in upper-class major fields of study and grades in Freshman prerequisites for these advanced divisions, then the predictions, which were developed from these Freshman criteria, will also show a low degree of relationship with upper-class majors.

CHAPTER VI

The Validity of the Forecasts in Upper-class Years

Before entering upon a discussion of the upper-class validity of the forecasts, it is necessary to make brief reference to the manner in which grade averages in the Junior and Senior Years were calculated. The marking systems in the three upper-class schools (Yale College, Sheffield Scientific School, and School of Engineering) are all based on different grading standards. In the latter two schools class marks are expressed in terms of letter grades, and in Yale College grades are expressed in numerical terms. It was necessary, therefore, to transmute all reported grades to a comparable basis. This was done in the following manner. Individual grade averages were transmuted to percentile scores for each school separately. Then, by means of a table of deviates of the normal curve for each permille of frequency, these percentiles were transmuted to standard scores with a mean of 75 and a standard deviation of 10.

It will be recalled from the discussion in Chapter IV that Freshmen are required to elect subjects in both the verbal and quantitative fields. By virtue of these distribution requirements, it was possible to establish a differential criterion against which to validate the differential predictions in the Freshman Year. When it comes to validating these predictions against achievement in the upper-class years, however, a new problem arises. After the Freshman Year, students begin to study in a field which they have chosen as their major concentration; and in the Junior and Senior Years the majority of the work is in the major field, and the rest is usually composed of closely allied subjects. Consequently, in a situation where course work is concentrated in a particular field of study, there is no possibility of establishing a differential criterion for Juniors and Seniors.

Without a measure of differential achievement in the upper-class years, the correlation procedure which was followed in

validating the differential forecasts in Freshman Year cannot be used to advantage. It would be possible, of course, to correlate the differences between verbal and quantitative predicted scores with upper-class grade averages. However, the coefficients thus obtained would have very little meaning, since there is no reason for supposing any particular degree of correspondence between these measures. For example, suppose one were to take the records of two students majoring in English, each of whom had a predicted difference of ten points in favor of verbal ability. Let it be assumed that this verbal idiosyncrasy is indicated for one student by a quantitative prediction of 65 and a verbal prediction of 75; and let these predicted scores for the other student be 85 and 95 respectively. In a situation such as this, which might easily occur in a considerable number of cases, there is no particular reason for believing that a student's achievement in his major field of work bears any relationship to the degree of difference between his verbal and quantitative forecasts. In other words, upper-class achievement must be judged with reference to individual *levels of ability*, as well as with respect to *differences* between these two abilities.

In order to analyze the validity of the predictions in the upper-class years, it is necessary, first, to determine which of the major fields of concentration are verbal and which are quantitative in essence. At the beginning of this thesis, English, History, French, German, Latin, and Spanish were defined as verbal subjects; and Mathematics, Chemistry, Physics, and Engineering Drawing were designated quantitative subjects. This differentiation was made partly on the basis of analysis and comparison of each type of material (Chapter I), and partly on the basis of certain experimental evidence (Table IV, page 19). In order to determine whether achievement in a particular field of major study may properly be called verbal or quantitative, it is necessary to correlate Junior and Senior Year averages with Freshman Year grade averages in each of the differentiated fields. These data are presented in Table IX.

In view of the relatively small degree of differentiation found between achievement in verbal and quantitative subjects of study

TABLE IX

DATA PERTAINING TO THE CORRELATION OF JUNIOR AND SENIOR YEAR GRADE AVERAGES IN VARIOUS FIELDS OF MAJOR STUDY AND THE AVERAGE OF GRADES IN FRESHMAN VERBAL AND QUANTITATIVE SUBJECTS

Major Field of Study (x variable)	Freshman Averages (y variable)	Mx	σx	My	σy	N	r	PE $_{(\text{diff. } r_1 - r_2)}$
Architecture and Art	Verbal Ave.	74.43	6.72	73.77	5.45	30	.44	
	Quant. Ave.	"	"	72.90	7.77	"	.51	.13
English	Verbal. Ave.	76.41	7.93	77.34	6.40	133	.65	
	Quant. Ave.	"	"	71.81	10.36	"	.50	.06
Foreign Languages	Verbal Ave.	77.48	8.43	77.72	6.44	51	.73	
	Quant. Ave.	"	"	70.34	11.00	"	.50	.08
History and H.A.L.	Verbal Ave.	74.32	8.61	75.36	6.78	212	.67	
	Quant. Ave.	"	"	70.61	10.35	"	.43	.05
Math. and Sci. in Yale Col.	Verbal Ave.	79.88	8.90	78.71	7.33	39	.47	
	Quant. Ave.	"	"	83.12	9.07	"	.55	.11
Medicine and Zoölogy	Verbal Ave.	74.02	8.86	76.52	5.13	101	.55	
	Quant. Ave.	"	"	74.01	9.55	"	.46	.07
Psychology and Philos.	Verbal Ave.	75.98	7.96	73.98	6.05	23	.76	
	Quant. Ave.	"	"	72.50	8.79	"	.34	.14
Social Sciences	Verbal Ave.	73.92	8.19	74.64	7.01	285	.68	
	Quant. Ave.	"	"	71.58	10.40	"	.54	.04
Applied Econ. Science	Verbal Ave.	74.70	6.96	73.57	6.26	140	.42	
	Quant. Ave.	"	"	71.13	8.82	"	.41	.07
Biol. and Plant Sci.	Verbal Ave.	81.11	6.95	76.46	6.41	49	.60	
	Quant. Ave.	"	"	77.19	8.31	"	.57	.09
Industrial Admin.	Verbal Ave.	74.30	7.05	74.12	6.94	101	.41	
	Quant. Ave.	"	"	77.71	8.47	"	.52	.07
Math. and Sci. in S.S.S.	Verbal Ave.	81.97	9.31	78.46	7.68	45	.62	
	Quant. Ave.	"	"	82.32	8.76	"	.61	.09
Chemical Engineering	Verbal Ave.	75.30	8.01	78.37	6.00	30	.46	
	Quant. Ave.	"	"	82.43	6.90	"	.55	.13
Civil Engineering	Verbal Ave.	76.98	8.16	74.00	5.87	28	.44	
	Quant. Ave.	"	"	79.14	6.53	"	.51	.14
Electrical Engineering	Verbal Ave.	76.14	10.26	74.32	5.46	22	.55	
	Quant. Ave.	"	"	82.23	5.30	"	.46	.15
Mechanical Engineering	Verbal Ave.	75.74	7.39	73.64	6.04	37	.22	
	Quant. Ave.	"	"	76.23	6.80	"	.48	.14

in Freshman Year, it is not surprising to discover that it is difficult to determine empirically which major fields of study in the upper-class years are verbal in nature and which are quantitative. A comparison of actual differences between the paired coefficients in the accompanying table and the probable errors of these differences indicates that History and H.A.L. (History, the Arts, and Letters) combined are the only majors which have a perfectly reliable difference. However, the chances are at least 95 in 100 that a true difference exists between the paired coefficients for English, Foreign Language, Psychology and Philosophy, and Social Science majors.

This analysis of correlation evidence indicates that the above-mentioned fields are verbal majors. No reliable differences in favor of quantitative majors are indicated by the coefficients in Table IX. This may be due in part to the fact that there are relatively few cases in most of the groups which would ordinarily be considered quantitative majors. (The $P.E._{(\mathrm{diff.}r_1-r_2)}$ is largely determined by the number of cases involved.) Since, however, the verbal and quantitative criteria in Freshman Year were not differentiated solely on the basis of empirical evidence, but also on the basis of a rational analysis of differences between subject-matter content and methods of dealing with each type of material, it seems logical at least to consider other factors as well as correlation evidence in attempting to differentiate upper-class majors. In so doing, the writer is aware that he is treading on rather weak ground. But since relatively few of these majors can be clearly differentiated on the basis of correlation evidence, only two courses of action are open: first, this study could be summarized and concluded at this point; or secondly, the rather meager evidence available might be further analyzed in order to see whether any positive conclusions could possibly be drawn. The latter course was deemed advisable.

In addition to the correlation evidence in Table IX, one might proceed on a basis of definition. Thus English, History, and Foreign Language majors would be designated verbal subjects, and Mathematics and Science in Yale College and S.S.S. (Sheffield Scientific School) would be called quantitative subjects in

accordance with the distinction made in Freshman Year. It is also possible to analyze the data in Table IX from the standpoint of differences between the means of the Freshman verbal and quantitative grade averages (My) for those students in each major field. It will be recalled from the discussion in Chapter V that a difference of four points or more between the Freshman verbal and quantitative criteria was considered significant.

Proceeding, then, on the basis of these three factors—definition, differences between the paired means, and differences between the paired coefficients—an attempt was made to distinguish between the verbal and quantitative major subjects of study. Naturally, this necessitated some compromise adjustments in instances where the various factors did not all point in the same direction. It was decided to experiment, first, with a combination of majors based upon the verbal and quantitative criteria established in Freshman Year. Thus English, Foreign Language, History, and H.A.L. were combined to form a verbal group of majors; and Mathematics and Science majors in Yale College and the Sheffield Scientific School and Chemical Engigineering were combined to form a quantitative group.

Data pertaining to the correlation of upper-class achievement in each of these combined fields of major study with the average of grades in Freshman verbal and quantitative subjects are presented in Table X. The resulting coefficients indicate a rather marked degree of association ($r=.68$) between achievement in the combined verbal majors and Freshman verbal averages, and a relatively low degree of correlation ($r=.46$) with Freshman quantitative averages. The difference between these *r*'s is statistically significant. On the other hand, achievement in the combined quantitative majors does not correlate particularly highly with either the Freshman verbal average ($r=.51$) or the Freshman quantitative average ($r=.55$). The difference between these coefficients is not statistically significant.

The number of possible permutations and combinations of upper-class majors could be greatly extended. Considering the general weakness of the data as a whole, however, it did not seem profitable to extend this portion of the study much further.

TABLE X

Data Pertaining to the Correlation of Junior and Senior Year Grade Averages in Various Combined Fields of Major Study with the Average of Grades in Freshman Verbal and Quantitative Subjects

Combined Major Fields of Study (x variable)	Freshman Averages (y variable)	Mx	σx	My	σy	N	r	$PE_{(diff. r_1 - r_2)}$
English, Foreign Language, History and H.A.L.	Verbal Ave.	75.43	8.45	76.33	6.69	396	.68	
	Quant. Ave.	"	"	70.98	10.46	"	.46	.03
English, Foreign Lang., Hist., H.A.L., Psych., Philos., and Social Science	Verbal Ave.	74.84	8.37	75.57	6.86	704	.68	
	Quant. Ave.	"	"	71.27	10.39	"	.48	.02
Math. and Science in Y.C. and S.S.S. and Chem. Eng.	Verbal Ave.	79.50	9.23	78.52	7.15	114	.51	
	Quant. Ave.	"	"	82.62	8.43	"	.55	.06
Math. and Science in Y.C. and S.S.S., Chem. Eng., Civ. Eng., Elect. Eng., Mech. Eng., Indust. Admin., Art and Arch.	Verbal Ave.	76.55	8.45	75.55	6.96	332	.48	
	Quant. Ave.	"	"	79.22	8.50	"	.54	.04
Biol. and Plant Science, Medicine and Zoölogy, and Applied Econ. Science	Verbal Ave.	75.55	8.08	75.09	6.10	290	.47	
	Quant. Ave.	"	"	73.16	9.27	"	.47	.04

Consequently, only three additional combinations were made. To the verbal majors, English, Foreign Language, History, and H.A.L. were added Psychology, Philosophy, and Social Science. The results obtained by correlating achievement in this larger group of majors with Freshman verbal and quantitative averages were almost identical with the data secured from a combination of English, Foreign Language, History, and H.A.L. majors alone. When Civil, Electrical, and Mechanical Engineering, Industrial Administration, Art and Architecture majors were combined with the Mathematics, Science, and Chemical Engineering group, the coefficients, means, and standard deviations were lowered, but the small degree of differentiation between each of these paired measures remained approximately the same. The data in Table IX did not seem to indicate that Biological and Plant Science, Medicine, Zoölogy, and Applied Economic Science were either clearly verbal or quantitative in character. Consequently these majors were combined in Table X to form an indeterminate or non-differentiated group. The lack of differentiation between the paired coefficients for this latter combination of majors appears to justify the grouping which was made.

The data presented in Tables IX and X were calculated for the purpose of determining, if possible, which of the upper-class major fields of study might properly be called verbal, and which quantitative, in character. From the analysis of these data it may be concluded that English, Foreign Language, History, History the Arts and Letters, Psychology, Philosophy, and the Social Sciences are clearly verbal majors. In the second place, Mathematics, the Sciences, Chemical, Civil, Electrical, and Mechanical Engineering, Industrial Administration, Art and Architecture majors appear to be more closely related to the quantitative criterion than to the verbal criterion. However, the degree of difference between the paired coefficients for the latter majors is small and relatively unreliable. And thirdly, the verbal or quantitative character of Biological and Plant Science, Medicine, Zoölogy, and Applied Economic Science majors is indeterminate.

The next step in this analysis is the validation of the verbal and quantitative predictions against achievement in upper-class curricula. The data pertaining to this problem, set forth in Table XI, are fairly comparable to the results obtained by correlating upper-class grade averages with the Freshman verbal and quantitative criteria (Table X). However, the degree of differentiation between the paired coefficients for the verbal majors, when correlated with the predictions (Table XI), is somewhat less reliable than the corresponding differences between the paired coefficients when these majors are correlated with the Freshman verbal and quantitative averages (Table X). This statement is likewise true with reference to the quantitative majors. Furthermore, the verbal majors do not correlate as highly with the verbal predictions (r=.56) as they do with the Freshman verbal criterion (r=.68), nor do the quantitative majors correlate as highly with the quantitative predictions (r=.50) as they do with the Freshman quantitative criterion (r=.54). This finding, in itself, is not startling, since the relationship between upper-class achievement and the predictions (which were originally developed from the Freshman criteria) is limited by the degree of association between these criteria and upper-class grade averages. The surprising fact is that grades in Freshman subjects of study, which presumably are prerequisite to major work in the Junior and Senior Years, are not more highly correlated with achievement in these advanced divisions.

This analysis of the validity of the predictions in the upper-class years must be carried a step further. In dealing with the validity of the predictions in Freshman Year (Chapter V) it was observed that most individuals evinced no marked differentiation between their verbal and quantitative forecasts; only about one-quarter of the cases (473 out of 1720) exhibited differences which were statistically significant. Furthermore, it was observed that mass correlations based on total class data tend to minimize the importance of large differences between the verbal and quantitative forecasts, since the latter exist only in a minority of cases. Yet this minority group, with predicted differences of eight or more points, represent the very cases for which a pro-

TABLE XI

DATA PERTAINING TO THE CORRELATION OF JUNIOR AND SENIOR YEAR GRADE AVERAGES IN VARIOUS COMBINED FIELDS OF MAJOR STUDY WITH THE VERBAL AND QUANTITATIVE PREDICTIONS

Combined Major Fields of Study (x variable)	Predictions (y variable)	Mx	σx	My	σy	N	r	$PE_{(diff.\,r_1 - r_2)}$
English, Foreign Language, History and H.A.L.	Verbal Pred.	75.43	8.45	75.69	7.09	396	.54	
	Quant. Pred.	"	"	71.60	9.97	"	.44	.04
English, Foreign Lang., Hist., H.A.L., Psych., Philos., and Social Science	Verbal Pred.	74.84	8.37	79.94	7.29	704	.56	
	Quant. Pred.	"	"	71.36	10.12	"	.45	.03
Math. and Science in Y.C. and S.S.S. and Chem. Eng.	Verbal Pred.	79.50	9.23	79.15	7.49	114	.52	
	Quant. Pred.	"	"	81.48	10.47	"	.55	.06
Math. and Science in Y.C. and S.S.S., Chem. Eng., Civ. Eng., Elect. Eng., Mech. Eng., Indust. Admin., Art and Arch.	Verbal Pred.	76.55	8.45	76.19	7.23	332	.48	
	Quant. Pred.	"	"	77.18	10.50	"	.50	.04
Biol. and Plant Science, Medicine and Zoölogy, and Applied Econ. Science	Verbal Pred.	75.55	8.08	74.20	7.08	290	.45	
	Quant. Pred.	"	"	72.53	10.27	"	.42	.05

gram of differential counseling might be of critical importance. It is necessary, therefore, to go beyond the mass correlations presented in Table XI in order to determine the validity of the forecasts for this critical minority group.

Of the 473 cases in the Classes of 1934, 1935, and 1936 who had large predicted differences, 91 failed to graduate, 81 graduated from major fields which could not be differentiated clearly as to their verbal or quantitative character, and 301 graduated from either verbal or quantitative major fields of study. This latter group represent the available cases upon which further analysis of the validity of the predictions in upper-class years may be based. Before examining the data presented in Table XII, it is important to understand the manner in which these 301 cases were grouped. Comparison of the verbal and quantitative predicted scores indicated that 118 of these cases appeared to have superior verbal ability, or a verbal idiosyncrasy. It is interesting to note that, of this number, 101 (86%) chose to enter a verbal curriculum (Group 1a), and only 17 entered a quantitative major (Group 1b). On the other hand, 183 cases appeared to have superior quantitative ability. Of this number, 75 entered a verbal major (Group 2a), and 108 entered a quantitative curriculum (Group 2b).

Data pertaining to the correlation of Junior and Senior Year grade averages, with the verbal and quantitative predictions for each of these groups, are set forth in Table XII. There are no statistically significant differences between the paired coefficients for any of these groups. If the predicted scores efficiently forecasted and differentiated achievement in verbal and quantitative majors, one would expect to find a relatively high degree of correlation between upper-class achievement and the verbal predictions, and a relatively low degree of correlation with the quantitative predictions for those students in Group 1a and Group 2a who entered a verbal major. And for those students in Group 1b and Group 2b who entered a quantitative major, one would expect to find achievement relatively highly correlated with the quantitative predictions, and the correlation with the verbal predictions somewhat lower. While the correlation coeffi-

TABLE XII

DATA PERTAINING TO THE CORRELATION OF JUNIOR AND SENIOR YEAR GRADE AVERAGES IN VARIOUS FIELDS OF MAJOR STUDY WITH THE VERBAL AND QUANTITATIVE PREDICTIONS

(Cases having a Predicted Difference of 8 points or more between their Verbal and Quantitative Forecasts)

(x variable) Grade averages in major fields of study for those cases whose predicted differences indicated:	(y variable) Predictions	Mx	σx	My	σy	N	r	$PE_{(diff.\ r_1 - r_2)}$
(1) *Verbal Idiosyncrasy* and who entered:								
(a) Verbal Major	Verbal Pred.	73.15	8.33	72.74	6.38	101	.40	
	Quant. Pred.	"	"	59.77	6.57	"	.35	.08
(b) Quantitative Major	Verbal Pred.	73.09	5.17	73.32	4.71	17	.22	
	Quant. Pred.	"	"	60.62	6.04	"	.31	.21
(2) *Quantitative Idiosyncrasy* and who entered:								
(a) Verbal Major	Verbal Pred.	76.66	8.07	77.27	6.61	75	.60	
	Quant. Pred.	"	"	84.21	6.76	"	.55	.07
(b) Quantitative Major	Verbal Pred.	79.67	8.91	77.83	7.18	108	.54	
	Quant. Pred.	"	"	85.83	8.12	"	.54	.07

cients in Table XII tend to verify these expectations, except in the case of Group 2b, the evidence is so meager that no positive conclusions with reference to these data seem justified.

Before concluding this analysis of the validity of the forecasts in the upper-class years for those individuals with large differences between their verbal and quantitative predicted scores, a comparison of the average performance (Mx) of the groups in Table XII might prove of interest. If no factors of self-selection were operating, one would expect those students with a verbal idiosyncrasy who entered a verbal major (Group 1a) to attain a higher average performance in the upper-class years than those students with a quantitative idiosyncrasy who entered a verbal major (Group 2a). The data indicate, however, that this is not the case. The mean of Group 2a is 3.51 points higher than the mean of Group 1a. This difference between the means is statistically significant. On the other hand, the average achievement of those students with a quantitative idiosyncrasy who entered a quantitative major (Group 2b) is higher by 6.58 points than the corresponding mean for those students with a verbal idiosyncrasy who entered a quantitative major (Group 1b). The difference between the mean upper-class performance of these two groups is also statistically significant. It is interesting to note that this latter difference, which is in the direction expected, contradicts the findings with reference to Groups 1a and 2a.

In order to understand the ambiguity or apparent contradiction of these findings, it is necessary to examine the assumption upon which the above analysis was made. The writer stated that one would expect average performance in the upper-class years to be higher for those students who entered a major field of study corresponding to their educational idiosyncrasy than for those students who entered a major contrary to the direction forecast. This statement rests, however, upon the postulate that the mean level of ability of the various groups in question is approximately equal. Unless this assumption is correct, any obtained difference between the average upper-class performance of these groups may be accounted for on the basis of actual differences between average levels of ability.

In judging the initial level of ability characteristic of Groups 1a, 2a, 1b, and 2b, any one of the following measures may be used: verbal or quantitative predicted scores, general predictions, grade averages in Freshman verbal or quantitative subjects of study, and Freshman general averages. Since the means and standard deviations of the distributions for these various measures and for the upper-class grade averages differ considerably, it was necessary to transmute them to a common basis for purposes of comparing the mean scores of each group. Of the 1720 students who had complete prematriculation and Freshman Year records, 1352 (79%) graduated. For these cases with altogether complete records, the means and standard deviations of the distributions of scores were found for each of the measures mentioned above. Then, for the 301 students with marked idiosyncrasies (Groups 1a, 1b, 2a, and 2b), the arithmetic means for each of the aforementioned measures of ability were transmuted to standard scores with the same mean and standard deviation as for the Freshman general averages. These data are set forth in Table XIII.

The average level of ability of Group 2a exceeds that of Group 1a, and the average level of ability of Group 2b is higher than that of Group 1b, as judged by all measures presented in Table XIII with the exception of the mean performance of these groups in Freshman verbal subjects of study. It is evident, therefore, that the differences between the upper-class performance of these groups, which were observed in Table XII, may be accounted for on the basis of self-selection, or actual differences between the levels of ability characteristic of these groups.

The data in Table XIII were calculated for the specific purpose of explaining an apparent contradiction between the mean upper-class performance of Group 1a as compared with Group 2a, and the performance of Group 1b as compared with Group 2b. Quite aside from this aim, these data happen to provide the basis for an interesting commentary on Yale's selective admission policy. It will be noted that the mean upper-class performance of those students whose predicted differences indicated a quantitative idiosyncrasy is higher than the average upper-class

TABLE XIII

DATA PERTAINING TO THE MEAN LEVEL OF PERFORMANCE CHARACTERIZING VARIOUS GROUPS OF STUDENTS WHO HAD A DIFFERENCE OF EIGHT POINTS OR MORE BETWEEN THEIR VERBAL AND QUANTITATIVE FORECASTS

(All measures transmuted to a comparable basis)

Group	Idiosyncrasy	Major Field	f	Verbal Prediction		Quant. Prediction		General Prediction		Freshman Verbal Average		Freshman Quant. Average		Freshman General Average		Average in Major Fields of Study	
				Mean	%ile	Mean	%ile	Mean	%ile	Mean	%ile	Mean	%ile	Mean	%ile	Mean	%ile
(1a)	V	V	101	73.33	*42*	67.05	*10*	70.61	*26*	74.98	*51*	69.58	*21*	72.71	*38*	73.61	*43*
(1b)	V	Q	17	73.86	*44*	67.58	*12*	71.56	*31*	76.56	*59*	74.19	*46*	73.94	*45*	73.56	*43*
(2a)	Q	V	75	77.46	*64*	82.44	*84*	79.17	*73*	76.33	*58*	78.35	*69*	77.15	*62*	76.38	*58*
(2b)	Q	Q	108	77.97	*67*	83.46	*87*	80.01	*75*	76.52	*59*	82.16	*83*	80.74	*78*	78.76	*71*

V=Verbal.
Q=Quantitative.

achievement of those students who possessed superior verbal ability. In other words, regardless of whether those students with a quantitative idiosyncrasy (Groups 2a and 2b) enter a verbal major or a quantitative major, their average performance is substantially higher than for those students with a verbal idiosyncrasy (Groups 1a and 1b). This evidence of the superiority of Groups 2a and 2b over Groups 1a and 1b in upper-class achievement is substantiated with remarkable consistency by the other measures of ability shown in Table XIII. Furthermore, 47 (22%) of the 217 students who possessed a verbal idiosyncrasy were dropped out of college for scholastic reasons, while only 12 (5%) of the 256 students with superior quantitative ability left college for that reason.

From these considerations it may be concluded that persons with superior quantitative ability are more rigidly selected at the time of entrance to Yale, than those students with superior verbal ability. This situation, however, is not the result of any conscious effort on the part of the Board of Admissions at Yale to be more lenient in the selection of students who possess a verbal idiosyncrasy than in the selection of students with superior quantitative ability. It just so happens that the verbal factors in secondary school curricula and in the entrance examination records so far exceed the quantitative factors, both in number and in the statistical weight attached to them, that a student with a quantitative idiosyncrasy competes at a disadvantage with other students for admission to Yale. The data in Table XIII indicate that those students with a quantitative idiosyncrasy who are admitted to Yale are not only above the average in verbal ability, but are extraordinarily outstanding in quantitative ability. In fact, it is because of this marked quantitative superiority that those students in Groups 2a and 2b exhibit any idiosyncrasy at all.

CHAPTER VII

CONCLUSION

The basic consideration underlying the work of this thesis has been the determination and validation of certain objective indices of individual ability for different types of curricula. It must be clear that a scientific guidance program cannot be realized until the various important type aptitudes for each student can be forecast on some uniform scale with reasonable accuracy. The attainment of this objective will be dependent not only upon the development of accurate predictive instruments, but also upon the establishment of reliable criteria against which to validate these forecasts. Such an ideal may be incapable of realization, but this, of course, can only be ascertained by trial.

The determination of the differential aptitudes of students for verbal and quantitative types of curricula is a matter of practical importance at Yale. Students desiring to study Mathematics or Science (quantitative subjects) generally enter the Sheffield Scientific School or the School of Engineering in the Sophomore Year; and those who are planning to major in English, Foreign Language, or Social Science (verbal subjects) enter Yale College. Because of variations in the prerequisite work demanded by each of the upper-class schools, important differentiation between individual programs of study actually begins in the Freshman Year at Yale. Consequently, it is advisable that the differential aptitudes of students be determined prior to entrance.

That there are real differences between the verbal and quantitative abilities of students is a major assumption of this study. It is conceivable that these more or less arbitrarily defined aptitudes are essentially alike, both of them being so largely dependent upon some central factor, such as "general intelligence," that no important difference exists between them. This alone would preclude the possibility of any differential guidance program.

In the absence of more objective achievement testing at Yale, grades in separate courses of study provide the only means for testing the assumption of disparate ability for verbal and quantitative types of subject matter. A study of the intercorrelation of these grade averages indicated the possible existence of unique factors, but the comparatively high correlation of +.53 between the average of Freshman Year grades in verbal and quantitative subjects of study also indicated the presence of a common or general factor to a fair degree.

As mentioned above, this relatively high degree of association between the differential criteria in the Freshman Year may be due, primarily, to the presence of some underlying common factor such as general intelligence. However, it seems reasonable to assume that a lack of greater differentiation between achievement in verbal and quantitative types of course material may also be attributed, in part at least, to the fact that students must successfully complete *all* their work in any one year before they are permitted to advance to the next rung in the educational ladder. This requirement is not, of course, just a peculiarity of the Freshman Year at Yale; it is characteristic of the curriculum demands at every school level. Regardless of what opinion one may hold as to the necessity or value of this requirement in the educational system at present, it should be clear that such a demand is an important motivational factor which tends to obscure rather than reveal individual educational idiosyncrasies. In other words, when successful completion of all subjects of study is set forth as a primary objective, a student will tend to approach his studies with the idea of not doing poorly in any one area of the curriculum rather than with the thought of doing particularly well in that type of work for which he may have unique ability of a relatively high order.

With these difficulties in mind, differences between the verbal and quantitative predicted scores, calculated from certain prematriculation data, were initially validated against corresponding differences between Freshman grade averages in these two fields. A criterion difference of four points or more was taken as a reasonable standard of validity in view of the tendency for

formal school requirements to reduce achievement in disparate types of study to a common level of performance and thereby minimize trait differences. Approximately one-quarter of the cases had differences of eight points or more between their verbal and quantitative predicted scores. For this minority group, it was clearly evident that the differential predictions developed in this study correctly forecasted criterion differences in the Freshman Year approximately two out of every three times.

In order to determine the ultimate amount of dependence which counselors might place upon these forecasts, they were also studied with reference to performance in upper-class major fields. In the Junior and Senior Years the majority of a student's work is in his major field, and the rest is usually composed of closely allied subjects. This fact eliminated the possibility of establishing a differential criterion in the upper-class years against which to validate the differential predictions. Consequently, it was necessary to utilize a different method in this portion of the study, from that employed in analyzing the validity of the forecasts in the Freshman Year.

Grade averages in each major field of study were correlated, first, with Freshman Year criterion grades in the two differentiated fields in order to determine which of the upper-class majors might properly be called verbal and which quantitative. Thus, for each major field two coefficients of correlation were secured. A comparison of the differences between each of these paired coefficients with the reliability of these differences indicated that the verbal or quantitative character of achievement in the various major fields could be clearly determined in only a relatively few instances. In addition to this lack of differentiation between the coefficients, the size of the correlations between Freshman Year performance and upper-class achievement in both the verbal and quantitative fields was surprisingly low. This latter finding is especially interesting in view of the tendency, on the part of the faculty of the three upper-class schools, to consider achievement in various Freshman subjects of study as valid indices of potential ability for more advanced work in upper-class major fields.

Lack of substantial relationship between upper-class performance and achievement in certain Freshman subjects of study, considered as prerequisite to more advanced work, was also found by the Department of Personnel Study at Yale in an investigation conducted in 1933. In that research, the results of which have not been published heretofore, the correlation of Senior Year grade averages in engineering subjects only, with Freshman grades in Mathematics, Chemistry, and Engineering Drawing yielded coefficients of +.38, +.35, and +.32, respectively. These data were based upon 149 cases in the Classes of 1932 and 1933. The reader should bear in mind that Mathematics, Chemistry, and Engineering Drawing are required of Freshmen by the School of Engineering at Yale. These studies are considered not only as valuable and necessary tool subjects for more advanced work, but achievement in them is also considered as a reliable forecast of a student's subsequent performance in the engineering curriculum. They have been characterized by faculty members as the " sieve " or " screen " by which engineering students are selected. Obviously, these Freshman subjects are not adequate for this latter purpose.

Several factors may account for the low correlations cited above. In the first place, these data may afford evidence of unequal rates of growth, from the Freshman to the Senior Year, on the part of the individuals composing the group. Secondly, ability to achieve success in Freshman Mathematics, Chemistry, and Engineering Drawing may be quite distinct from that ability underlying achievement in engineering subjects of study. And thirdly, these low correlations may be accounted for on the basis of the unreliability of course grades. While it is quite possible that each of these factors contributes in some measure to the low correlations between engineering grades in the Senior Year and Freshman prerequisites to such advanced work, the writer is inclined to believe that the unreliability of classroom marks is the primary causal factor.

This contention appears to be verified by the correlation of +.72 found between grade averages in Junior engineering subjects and achievement in Senior engineering work for 165 cases

in the Classes of 1932, 1933, and 1934 at Yale. While a correlation of +.72 is relatively high as compared with the relationship found between grade averages in Senior engineering subjects and Freshman grades in Mathematics, Chemistry, and Engineering Drawing, it is much lower than one would expect it to be. The subjects comprising the engineering curricula in the Junior and Senior Years at Yale are closely related to each other; the content of these courses represents a relatively compact, sequential, and highly homogeneous type of material. Furthermore, it is not likely that the academic standing of students, relative to each other, would change very much between the Junior and Senior Years. If these presumptions are correct, then it is reasonable to assume that the lack of substantially higher correlation between sequential course work in engineering at Yale may be attributed, in no small part, to the unreliability of the measures of achievement.

Although the data which have just been introduced are not closely related to the work of this thesis, nevertheless, they afford additional evidence of lack of reliability or stability of criterion grades at Yale, which tends to support and verify the findings of this research. The importance of having reliable criteria against which to validate one's forecasts cannot be emphasized too much. The relatively low correlations which were found between grades in upper-class major fields and grades in Freshman Year subjects of study places a definite limitation on the development of valid predictive instruments of the type dealt with in this study.

It was not surprising to find, therefore, that the aptitude forecasts, which were developed from prematriculation data and originally designed to predict grade averages in Freshman verbal and quantitative subjects, failed to differentiate between achievement in verbal and quantitative major fields of study in the upper-class years. This lack of differentiating power on the part of the predictions was noted both with reference to the total number of cases and with reference to that minority group which had large differences of eight points or more between their verbal and quantitative forecasts.

It is necessary to conclude, therefore, that the major findings of this research are negative. In other words, the objective indices of differential aptitude developed in this study are not sufficiently valid to warrant their use in guiding and counseling students at Yale with reference to their choice of a major field of study. Whereas these forecasts appear to have some value in estimating differential achievement in the Freshman Year, there is practically no evidence that they differentiate between the upper-class performance of students in verbal and quantitative fields of study.

On the basis of the evidence presented in this dissertation, two important recommendations can be made. In the first place, there is a rather strong indication that current measures of achievement at Yale are relatively unreliable indices of student performance. As a means of improving the reliability of course grades, the writer urges the substitution of well-standardized objective achievement tests for those measures now in use. The lack of reliable criteria is a serious limitation upon the further development of valid forecasts of achievement.

Secondly, the evidence presented in the last chapter clearly indicates that students who possess superior quantitative ability represent a more highly selected group at the time of admittance to Yale than those students who have a verbal idiosyncrasy. This situation, which is not the result of any conscious effort on the part of the Board of Admissions to legislate against those students who possess a quantitative idiosyncrasy, is nevertheless unfortunate. The writer strongly recommends that serious consideration be given to the possibility of improving the admission policy of Yale University by giving greater weight to the quantitative factors in the pre-entrance records of students.

Bibliography

1. Bennett, George K.: *Differential Aptitude Prognosis* (unpublished dissertation), Yale University, 1935.
2. College Entrance Examination Board: *Twenty-seventh Annual Report of the Secretary and Second Annual Report of the Chairman of the Commission on Scholastic Aptitude Tests*, New York, 1927.
3. College Entrance Examination Board: *Twenty-eighth Annual Report of the Secretary and Third Annual Report of the Chairman of the Commission on Scholastic Aptitude Tests*, New York, 1928.

4. College Entrance Examination Board: *Twenty-ninth Annual Report of the Secretary and Fourth Annual Report of the Chairman of the Commission on Scholastic Aptitude Tests,* New York, 1929.
5. College Entrance Examination Board: *Thirtieth Annual Report of the Secretary and Fifth Annual Report of the Commission on Scholastic Aptitude Tests,* New York, 1930.
6. College Entrance Examination Board: *Thirty-third Annual Report of the Secretary and Eighth Annual Report of the Commission on Scholastic Aptitude Tests,* New York, 1933.
7. College Entrance Examination Board: *Thirty-fourth Annual Report of the Secretary and Ninth Annual Report of the Commission on Scholastic Aptitude Tests,* New York, 1934.
8. College Entrance Examination Board: *Thirty-fifth Annual Report of the Secretary and Tenth Annual Report of the Commission on Scholastic Aptitude Tests,* New York, 1935.
9. Crawford, Albert B.: Aptitude Testing in Personnel Procedure, *Bulletin of the American Association of Collegiate Registrars,* 1934, N.S. **9**, 293–309.
10. Crawford, Albert B.: Forecasting Certain College Aptitudes, *Person. J.,* 1932, **11**, 160–168.
11. Crawford, Albert B.: Individual Differences in Educational Guidance, *Psychol. Monog.,* 1936, **47**, 148–172.
12. Crawford, Albert B.: *Report of the Director of the Department of Personnel Study and Bureau of Appointments,* Bulletin Yale Univ. (Supplement) for 1931–1932, 18–19.
13. Crawford, Albert B.: Some Criticisms of Current Practice in Educational Measurements, *Harvard Teachers Record,* 1933, **3**, No. 2.
14. Crawford, Albert B., and Burnham, Paul S.: Entrance Examinations and College Achievement, *Sch. & Soc.,* 1932, **36**, 344–352, 378–384.
15. Garrett, Henry E.: *Statistics in Psychology and Education,* Longmans, New York, 1926.
16. Holzinger, Karl J.: *Statistical Methods for Students in Education,* Ginn, Boston, 1928.
17. Hull, Clark L.: *Aptitude Testing,* World Book, Yonkers, 1928.
18. Hull, Clark L.: Variability in Amount of Different Traits Possessed by the Individual, *Jour. Educ. Psychol.,* 1927, **18**, 97–106.
19. Hull, Clark L., and Limp, Charles E.: The Differentiation of the Aptitudes of an Individual by Means of Test Batteries, *Jour. Educ. Psychol.,* 1925, **16**, 73–88.
20. Kelley, Truman L.: *Educational Guidance,* Columbia Univ., 1914.
21. Kelley, Truman L.: *Interpretation of Educational Measurements,* World Book, Yonkers, 1927.
22. Lee, J. Murray, and Segel, David: The Utilization of Data from Simple or Direct Prediction in the Development of Regression Equations for Differential Prediction, *Jour. Educ. Psychol.,* 1933, **24**, 550–554.
23. Limp, Charles E.: The Use of the Regression Equation in Determining the Aptitudes of an Individual, *Jour. Educ. Psychol.,* 1925, **16**, 414–418.
24. Segel, David: Differential Prediction of Ability as Represented by College Subject Groups, *Jour. Educ. Res.,* 1932, **25**, 14–26, 93–98.
25. Segel, David: Differential Prediction of Scholastic Success, *Sch. & Soc.,* 1934, **39**, 91–96.

26. SEGEL, DAVID, and GERBERICH, J. R.: Differential College Achievement Predicted by the American Council Psychological Examination, *Jour. Appl. Psychol.*, 1933, 17, 637–645.
27. THORNDIKE, EDWARD L.: Intelligence and Its Uses, *Harper's*, 1920, **140**, 227–235.
28. THORNDIKE, EDWARD L.: On the Organization of Intellect, *Psychol. Rev.*, 1921, **28**, 141–151.
29. WOOD, BEN D.: *Measurement in Higher Education*, World Book, Yonkers, 1923.